Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications 385 (2004) 63–72

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Joint zero sets and ranges of several Hermitian forms over complex and quaternionic scalars

P. Binding [a,*], A. Markus [b]

[a]*Department of Mathematics and Statistics, University of Calgary, Calgary, AB, Canada T2N 1N4*
[b]*Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel*

Received 2 January 2003; accepted 18 March 2003

Submitted by L. Rodman

To our friend and colleague Peter Lancaster on his 75th birthday

Abstract

Let $h_j : V \to \mathbb{R}$, $j = 1, \ldots, n$, be Hermitian forms on a inner product space over $\mathbb{F} = \mathbb{C}$ or $\mathbb{H}$, and let $\mathbf{h} : V \to \mathbb{R}^n$ have jth component h_j. We study path connectedness of the joint zero set $\mathbf{h}^{-1}(\mathbf{0}) \cap V_1$ and convexity of the joint range $\mathbf{h}(V_1)$ for various values of n, where V_1 is the unit sphere of V.

AMS classification: 15A33; 15A60; 15A63; 47A12

Keywords: Quaternionic inner product space; Hermitian form; Joint zero set; Joint range; Hausdorff–Toeplitz theorem

1. Introduction

It is generally accepted that the study of numerical ranges of operators on complex inner product spaces started with the classic papers of Toeplitz [20] and Hausdorff [13]. It is not always appreciated, however, that the study of joint ranges of a general number n of Hermitian forms also started with these papers. There have been many investigations of the case $n = 2$ for complex scalars, corresponding to the above numerical ranges, but joint ranges for further values of n and/or scalars in $\mathbb{R}$ or $\mathbb{H}$ (the

* Corresponding author. Tel.: +403-220-6328.
E-mail addresses: binding@ucalgary.ca (P. Binding), markus@math.bgu.ac.il (A. Markus).

0024-3795/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0024-3795(03)00493-2

quarternions) have also been discussed. See, for example, Brickman [8] ($\mathbb{R}$), several papers of Au-Yeung and collaborators e.g. [2,4,5] ($\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$), Binding [6] ($\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$), Horn and Johnson [14, pp. 85, 86, 88] ($\mathbb{R}$ and $\mathbb{C}$) and Lyubich and Markus [17] ($\mathbb{C}$). Applications of joint ranges (which are called also vectorial ranges, n-form numerical ranges and n-dimensional fields of values) have been made to stability of feedback control (see e.g. [9–11] and references there) and to multiparameter spectral theory [1,7].

Hausdorff [13] obtained the first proof of the convexity of numerical ranges (i.e., of joint ranges of two Hermitian forms) as a corollary of the connectedness of the zero set of a Hermitian form. This approach was continued in [17] where the connectedness of the joint zero set of two Hermitian forms was proved, and as a corollary a new proof of the convexity of the joint range of three Hermitian forms was obtained. Lyubich [16] used the connectedness result from [17] to prove a general theorem on the separation of roots of matrix and operator polynomials. Partial results in this direction were obtained earlier in [18] with the help of the theorem on convexity of joint ranges for three Hermitian forms [4,6].

To fix ideas, let V be an inner product space over $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ (see e.g. [15, Chapter 13] for the case $\mathbb{F} = \mathbb{H}$). Noting that $\mathbb{H}$ is noncommutative, we multiply by scalars on the left, so V is then a left vector space. As usual, we define the norm in V by $\|x\| = \langle \mathbf{x}, \mathbf{x}\rangle^{1/2}$ where $\langle \mathbf{x}, \mathbf{y}\rangle$ is the inner product. The set of $\mathbf{x} \in V$ with positive norm, i.e., $\mathbf{x} \neq \mathbf{0}$ (resp. norm 1, i.e., the unit sphere) is denoted by V_+ (resp. V_1). A form $\sigma : V \times V \to \mathbb{F}$ is *sesquilinear* if $\sigma(\mathbf{x}, \mathbf{y})$ is linear in $\mathbf{x}$ for each fixed $\mathbf{y}$ and $\sigma(\mathbf{y}, \mathbf{x}) = \overline{\sigma(\mathbf{x}, \mathbf{y})}$. A form $h_\ell : V \to \mathbb{R}$ is *Hermitian* if $h_\ell(\mathbf{x}) = \sigma_\ell(\mathbf{x}, \mathbf{x})$ for some sesquilinear σ_ℓ. If $\mathbf{h} : V \to \mathbb{R}^n$ has Hermitian components h_ℓ, we call $W_1 = \mathbf{h}(V_1)$ the *joint range* of the h_ℓ. (This term is used for $W_+ = \mathbf{h}(V_+)$ in [12].) We call $Z_1 = \mathbf{h}^{-1}(\mathbf{0}) \cap V_1 = \{\mathbf{x} \in V_1 : \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ the *joint zero set* of the h_ℓ. We shall also consider $Z = \mathbf{h}^{-1}(\mathbf{0})$ and $Z_+ = \mathbf{h}^{-1}(\mathbf{0}) \cap V_+ = \{\mathbf{x} \neq 0 : \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$.

In Sections 2 and 3 we study path connectedness of the joint zero set. This is simple for Z, but more interesting for Z_1 (and for Z_+, where the property is equivalent). Let d be the dimension of $\mathbb{F}$, viewed as a real vector space, so $d = 2$ for $\mathbb{F} = \mathbb{C}$ and $d = 4$ for $\mathbb{F} = \mathbb{H}$. In Section 2 we show for $\mathbb{F} = \mathbb{C}$ or $\mathbb{H}$ that Z_+ is path connected if $n < d$ and in Section 3 we improve this to $n \leqslant d$ provided $\dim V > 2$. These results are known for $\mathbb{F} = \mathbb{C}$ [17] but our proof is simpler, and it seems to us that the proof in [17] cannot easily be extended to $\mathbb{F} = \mathbb{H}$.

In Section 4 we consider convexity of the joint range W_1. It will be seen that this implies convexity of W_+, but not conversely—see Corollary 4.3 and Remark 5.6(b). In Lemma 4.1 we develop the approach of [13,17] and claim that every result about connectedness of joint zero sets implies a corresponding result about convexity of joint ranges, and in this passage the number of Hermitian forms may be increased by one. Hence the results of Sections 2 and 3 imply that W_1 is convex for $n \leqslant d$, and if $\dim V > 2$ even for $n \leqslant d + 1$. These convexity results are known ([4,6] for $\mathbb{C}$, [6] for $\mathbb{H}$) but again our proof is more elementary than those in print. In Section 5 we give examples showing optimality of our results.

The case $\dim V = 1$ is special, and may be treated via the identity

$$h_\ell(\alpha\mathbf{x}) = |\alpha|^2 h_\ell(\mathbf{x}) \tag{1.1}$$

for any Hermitian form h_ℓ. Indeed, we have the obvious

Theorem 1.1. *If V has basis $\mathbf{x}$, then Z_1 is empty if $\mathbf{h}(\mathbf{x}) \neq \mathbf{0}$, or isomorphic to the unit sphere of $\mathbb{F}$ if $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, and hence is path connected if $\mathbb{F} = \mathbb{C}$ or $\mathbb{H}$. Moreover W_1 is a singleton, and hence is convex.*

We have excluded $\mathbb{F} = \mathbb{R}$ so far, even though most of our arguments apply. The reason is that the unit sphere of $\mathbb{R}$ is not connected, and therefore Z_1 need not be so (even if $\dim V = 1$, see Theorem 1.1). The case $\mathbb{F} = \mathbb{R}$ is discussed in [8,17]. Most of our reasoning also applies if $\mathbb{F}$ is the division algebra of octonions, but again a critical property ((1.1) in this case) is missing, and we intend to pursue this situation elsewhere.

Finally let us return to numerical ranges, this time over $\mathbb{H}$. We call a linear operator $A : V \to V$ Hermitian if $\mathbf{x} \to \langle A\mathbf{x}, \mathbf{x}\rangle$ is a Hermitian form. If a linear operator B admits a Hermitian decomposition $B_1 + \mathrm{i}B_2 + \mathrm{j}B_3 + \mathrm{k}B_4$ (see Section 2 for notation), i.e., where each B_ℓ is Hermitian, then the numerical range $W(B) = \{\langle B\mathbf{x}, \mathbf{x}\rangle : \mathbf{x} \in V_1\}$ is isomorphic to the joint range W_1 of four Hermitian forms, and is thus convex by our results. On the other hand, $W(B)$ need not be convex (even if $\dim V = 1$, see Section 5 for an example), so linear operators over $\mathbb{H}$ do not admit Hermitian decomposition in general. For further discussion of numerical ranges over $\mathbb{H}$, we refer to [3,14, pp. 86–88;19] and the references there.

2. Preliminaries

From now on we assume $F = \mathbb{C}$ or $\mathbb{H}$, although as indicated earlier, many of our arguments hold more generally.

We start by developing some properties of quaternions that we shall need. Each $a \in \mathbb{H}$ can be written $b + c\mathrm{j}$ where $b, c \in \mathbb{C}$ and $\mathrm{j}^2 = -1$. Thus

$$a = a_1 + \mathrm{i}a_2 + \mathrm{j}a_3 + \mathrm{k}a_4$$

for

$$\mathbf{a} = (a_1, a_2, a_3, a_4) \in \mathbb{R}^4 \tag{2.1}$$

and $\mathrm{k} = \mathrm{ij}$. Noncommutativity comes from the relation $\mathrm{ji} = -\mathrm{k}$, and conjugacy is defined by $\bar{a} = a_1 - \mathrm{i}a_2 - \mathrm{j}a_3 - \mathrm{k}a_4$. Further constructions are extended from $\mathbb{C}$ via $\operatorname{Re} a = a_1$ and $|a|^2 = a\bar{a} = a_1^2 + \cdots + a_4^2$, and we note that $\overline{ab} = \bar{b}\bar{a}$.

We shall need the following result on simultaneous equations.

Lemma 2.1. *If $p, q, r, s \in \mathbb{H}$, the following are equivalent*:

(a) $\mathbf{p}, \mathbf{q}, \mathbf{r}$ *and* $\mathbf{s}$ *are linearly independent in* $\mathbb{R}^4$*—see* (2.1),
(b) *there exists nonzero $a \in \mathbb{H}$ such that*

$$\mathrm{Re}(ap) = \mathrm{Re}(aq) = \mathrm{Re}(ar) = \mathrm{Re}(as) = 0, \tag{2.2}$$

(c) *for some $t, u, v, w \in \mathbb{R}$, the system*

$$\mathrm{Re}(ap) = t,\ \mathrm{Re}(aq) = u,\ \mathrm{Re}(ar) = v,\ \mathrm{Re}(as) = w \tag{2.3}$$

has no solution for $a \in \mathbb{H}$.

Proof. (2.2) and (2.3) are linear equations in a_1, a_2, a_3 and a_4 with coefficient matrix

$$M = \begin{bmatrix} p_1 & -p_2 & -p_3 & -p_4 \\ q_1 & -q_2 & -q_3 & -q_4 \\ r_1 & -r_2 & -r_3 & -r_4 \\ s_1 & -s_2 & -s_3 & -s_4 \end{bmatrix}.$$

Evidently (a)–(c) are all equivalent to $\det M = 0$. □

Remark 2.2. Similarly if $p, q \in \mathbb{C}$ then the following are equivalent:

(a) $(\mathrm{Re}\, p, \mathrm{Im}\, p)$ and $(\mathrm{Re}\, q, \mathrm{Im}\, q)$ are linearly dependent in $\mathbb{R}^2$,
(b) there exists nonzero $a \in \mathbb{C}$ so that

$$\mathrm{Re}(ap) = \mathrm{Re}(aq) = 0, \tag{2.4}$$

(c) for some $t, u \in \mathbb{R}$, the system $\mathrm{Re}(ap) = t$, $\mathrm{Re}(aq) = u$ is insoluble for $a \in \mathbb{C}$.

In discussing path connectedness of zero sets, we shall use the notation $\mathbf{x} \sim \mathbf{y}$ to denote the existence of a path joining $\mathbf{x}$ to $\mathbf{y}$ within Z_+, so $\mathbf{f}(0) = \mathbf{x}$, $\mathbf{f}(1) = \mathbf{y}$ and $\mathbf{f} : [0, 1] \to Z_+$ is continuous. It will be unnecessary for us to have separate notation for Z and Z_1, in view of the following.

Lemma 2.3
(a) *If $\mathbf{x}, \mathbf{y} \in Z_+$ are linearly dependent then $\mathbf{x} \sim \mathbf{y}$.*
(b) *If Z_1 is path connected so is Z_+, and conversely.*
(c) *Z is always path connected.*

Proof. (a) Let $\mathbf{x}, \mathbf{y} \in Z_+$ be linearly dependent, say $\mathbf{x} = p\mathbf{y}$ where $0 \neq p \in \mathbb{F}$. Since $\mathbb{F}_+ := \mathbb{F} \setminus \{0\}$ is path connected, we can join 1 to p within $\mathbb{F}_+$, and hence we can join $1\mathbf{y}$ to $p\mathbf{y}$ within Z_+.

(b) This is proved for $\mathbb{F} = \mathbb{C}$ in [17, Lemma 1.2] and the proof for $\mathbb{F} = \mathbb{H}$ is identical.

(c) By (1.2) we can join any $\mathbf{x} \in Z$ to $\mathbf{0} \in Z$ using the path $\mathbf{f}(t) = t\mathbf{x}$, $0 \leqslant t \leqslant 1$. Thus any $\mathbf{x}, \mathbf{y} \in Z$ can be joined via $\mathbf{0}$. □

The next step is to link equations like (2.2) to the equivalence relation $\sim$. Recall that the h_ℓ are generated by the sesquilinear forms σ_ℓ, $\ell = 1, \ldots, n$.

Lemma 2.4. *If* $\mathbf{x}, \mathbf{y} \in Z_+$ *and if there exists nonzero* $a \in \mathbb{F}$ *so that*

$$\operatorname{Re}(a\sigma_\ell(\mathbf{x}, \mathbf{y})) = 0, \quad \ell = 1, \ldots, n, \tag{2.5}$$

then $\mathbf{x} \sim \mathbf{y}$.

Proof. If $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, then $\mathbf{x} \sim \mathbf{y}$ by Lemma 2.3(a), so assume $\mathbf{x}$ and $\mathbf{y}$ to be linearly independent, and consider

$$\mathbf{f} : [0, 1] \to V : t \to ta\mathbf{x} + (1 - t)\mathbf{y}.$$

Evidently $\mathbf{f}(t) \neq \mathbf{0}$, and

$$h_\ell(\mathbf{f}(t)) = t^2 h_\ell(a\mathbf{x}) + 2t(1 - t)\operatorname{Re}(a\sigma_\ell(\mathbf{x}, \mathbf{y})) + (1 - t)^2 h_\ell(\mathbf{y}) = 0$$

by (1.1) and (2.5). Thus $a\mathbf{x} \sim \mathbf{y}$, and $\mathbf{x} \sim a\mathbf{x}$ by Lemma 2.3(a). $\square$

We are now in a position to establish a preliminary result on path connectedness of Z_+. Note that for $\mathbb{F} = \mathbb{C}$ this result was proved already by Hausdorff [13].

Recall that d is the dimension of $\mathbb{F}$ considered as a vector space over $\mathbb{R}$.

Theorem 2.5. *If* $n < d$ *then* $Z_+ = \{\mathbf{x} \neq \mathbf{0} : h_\ell(\mathbf{x}) = 0, \ell = 1, \ldots, n\}$ *is path connected*.

Proof. Appending zero forms if necessary, we can assume that $n = d - 1$. Let $\mathbf{x}, \mathbf{y} \in Z_+$, $p = \sigma_1(\mathbf{x}, \mathbf{y})$ and consider the Eq. (2.4) with $q = 0$ if $\mathbb{F} = \mathbb{C}$ (resp. (2.2) with $q = \sigma_2(\mathbf{x}, \mathbf{y})$, $r = \sigma_3(\mathbf{x}, \mathbf{y})$, $s = 0$ if $\mathbb{F} = \mathbb{H}$). Then Remark 2.2 (resp. Lemma 2.1) show that a nonzero $a \in \mathbb{F}$ exists satisfying (2.5). The result now follows from Lemma 2.4. $\square$

3. Path connectedness

In order to prove the main results of this section we need two lemmas. The first is a simple consequence of the results of Section 2.

Lemma 3.1. *Let* $\mathbf{x}, \mathbf{y} \in Z_+$. *If* $\sigma_\ell(\mathbf{x}, \mathbf{y})$, $\ell = 1, \ldots, d$, *are linearly dependent* (*viewed as elements of* $\mathbb{R}^d$) *then* $\mathbf{x} \sim \mathbf{y}$.

Proof. By Lemma 2.1 (resp. Remark 2.2), the Eq. (2.2) (resp. (2.4)) have a nonzero solution $a \in \mathbb{F}$, and Lemma 2.4 completes the argument. $\square$

The second lemma is a parametrized version of the fact that $d-1$ homogeneous linear equations in d reals have a nontrivial solution (already used to prove Theorem 2.5). While it may be known, we have not seen it in print.

Lemma 3.2. *If $A(t)$ is a $(d-1)\times d$ matrix whose entries are real polynomials in t then the linear system $A(t)\mathbf{g}=\mathbf{0}$ has a solution $\mathbf{g}=\mathbf{g}(t)$ whose entries are real polynomials satisfying $\sum_{\ell=1}^{d} g_\ell(t)^2 \neq 0$ for all real t.*

Proof. If $A(t)\equiv 0$ then we set $g_\ell(t)\equiv 1$ for $\ell=1,\dots,d$. Otherwise let m ($\geqslant 1$) be the maximal size of a subdeterminant of $A(t)$ which does not vanish identically. We can assume without loss that this corresponds to the first m rows and columns of $A(t)$.

Consider the determinant $D(t)$ given by the first $m+1$ rows and columns of $A(t)$ (if $m=d-1$ we append a row of zeros). Let $\gamma_\ell(t)$ be the $(m+1,\ell)$ cofactor of $D(t)$ for $1\leqslant \ell\leqslant m+1$ and $\gamma_\ell(t)=0$ for $m+1<\ell\leqslant d$. It is easy to see that $A(t)\gamma(t)=\mathbf{0}$ (note that the $(m+\ell)$th row of $A(t)$ for $\ell\geqslant 1$ is linearly dependent on the first m rows) and $\nu(t):=\sum_{\ell=1}^{d}\gamma_\ell(t)^2\geqslant \gamma_{m+1}(t)^2\not\equiv 0$. Thus $\nu(t)$ is a nonzero nonnegative polynomial with real roots t_α of multiplicity $2m_\alpha$, say. Defining

$$g_\ell(t)=\gamma_\ell(t)\prod_\alpha (t-t_\alpha)^{-m_\alpha}$$

we see that the polynomials g_ℓ have no common real roots, and hence $\sum_{\ell=1}^{d} g_\ell(t)^2\neq 0$ for all real t. □

We are now ready to improve Theorem 2.5 by allowing $n=d$, provided $\dim V>2$.

Theorem 3.3. *If $\dim V>2$ and $n\leqslant d$ then Z_+ is path connected.*

Proof. In view of Theorem 2.5, we can assume that $n=d$.

Let $\mathbf{x},\mathbf{y}\in Z_+$. We wish to show $\mathbf{x}\sim\mathbf{y}$, and by Lemma 2.3(a) we may assume that $\mathbf{x}$, $\mathbf{y}$ and some vector $\mathbf{z}$ form a basis of a three dimensional subspace $L\subset V$. We claim that either $\mathbf{x}\sim\mathbf{y}$ or there exists nonzero $\mathbf{w}=a\mathbf{x}+b\mathbf{y}+\mathbf{z}$ satisfying

$$\sigma_d(\mathbf{w},\mathbf{x})=\sigma_d(\mathbf{w},\mathbf{y})=0. \tag{3.1}$$

If $\sigma_d(\mathbf{x},\mathbf{y})=0$ then Lemma 3.1 gives $\mathbf{x}\sim\mathbf{y}$ so assume $\sigma_d(\mathbf{x},\mathbf{y})\neq 0$. Thus (3.1) holds provided $a=-\sigma_d(\mathbf{z},\mathbf{y})/\sigma_d(\mathbf{x},\mathbf{y})$ and $b=-\sigma_d(\mathbf{z},\mathbf{x})/\sigma_d(\mathbf{x},\mathbf{y})$, and our claim is established. Note that

$$\mathbf{x},\mathbf{y},\mathbf{w}\text{ also form a basis of } L. \tag{3.2}$$

Suppose first that $\mathbf{w}\in Z_+$. Then (3.1) and Lemma 3.1 show that $\mathbf{w}\sim\mathbf{x}$ and similarly $\mathbf{w}\sim\mathbf{y}$, so $\mathbf{x}\sim\mathbf{y}$. Thus we may assume that $\mathbf{w}\notin Z_+$, i.e.

$$\mathbf{h}(\mathbf{w})\neq\mathbf{0}. \tag{3.3}$$

By Lemma 3.1 again we may assume that $\sigma_\ell(\mathbf{x},\mathbf{y})$ for $\ell = 1,\dots,d$ are linearly independent (viewed as elements of $\mathbb{R}^d$). Thus, Lemma 2.1 and Remark 2.2 show that the system

$$\operatorname{Re}(a\sigma_\ell(\mathbf{x},\mathbf{y})) = -\tfrac{1}{2}h_\ell(\mathbf{w}), \quad \ell = 1,\dots,d \tag{3.4}$$

has a solution $a \in \mathbb{F}$ and by (3.3) we have $a \neq 0$.

Writing $\widetilde{\mathbf{x}} = a\mathbf{x}$, we have $\mathbf{x} \sim \widetilde{\mathbf{x}}$ by Lemma 2.3(a), so it will suffice to prove $\widetilde{\mathbf{x}} \sim \mathbf{y}$, and to this end we consider

$$\mathbf{f}(t) = t\widetilde{\mathbf{x}} + (1-t)\mathbf{y} + g(t)\mathbf{w}, \tag{3.5}$$

where $g : [0,1] \to \mathbb{F}$ and $g(0) = g(1) = 0$. By (3.2), $\mathbf{f}(t) \neq \mathbf{0}$, so it remains to show that $\mathbf{h}(\mathbf{f}(t)) = \mathbf{0}$, for all $t \in [0,1]$. Using $\mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{y}) = \mathbf{0}$ we thus require

$$\begin{aligned}&2t\operatorname{Re}(g(t)\sigma_\ell(\mathbf{w},\widetilde{\mathbf{x}})) + 2(1-t)\operatorname{Re}(g(t)\sigma_\ell(\mathbf{w},\mathbf{y}))\\ &\quad + 2t(1-t)\operatorname{Re}\sigma_\ell(\widetilde{\mathbf{x}},\mathbf{y}) + |g(t)|^2 h_\ell(\mathbf{w}) = 0\end{aligned}$$

for $\ell = 1,\dots,d$. Setting

$$|g(t)| = t^{1/2}(1-t)^{1/2} \tag{3.6}$$

and recalling (3.4), we therefore need

$$t\operatorname{Re}(g(t)\sigma_\ell(\mathbf{w},\widetilde{\mathbf{x}})) + (1-t)\operatorname{Re}(g(t)\sigma_\ell(\mathbf{w},\mathbf{y})) = 0, \quad \ell = 1,\dots,d-1 \tag{3.7}$$

since by (3.1) the corresponding equation for $\ell = d$ is automatically satisfied.

Writing $\mathbf{g}(t)$ for the element of $\mathbb{R}^d$ corresponding to $g(t) \in \mathbb{F}$, we see that (3.7) is a set of equations to which Lemma 3.2 applies. Viewing the corresponding solution as an element of $\mathbb{F}$ and scaling it to satisfy (3.6), we can thus construct $\mathbf{f}(t)$ as in (3.5), and the proof is complete. □

4. Convexity of joint ranges

In order to reduce our discussion to the results of the previous sections, we shall use the following

Lemma 4.1. *Suppose* $\dim V < \infty$ *and* $n \in \mathbb{N}$. *If the set* Z_1 *is connected for* $n-1$ *arbitrary Hermitian forms over* V *then the set* W_1 *is convex for* n *arbitrary Hermitian forms over* V.

Proof. Let $\mathbf{u},\mathbf{v} \in W_1 \subset \mathbb{R}^n$ and $\mathbf{w} = t\mathbf{u} + (1-t)\mathbf{v}$ where $0 < t < 1$. Choose $\mathbf{p} \in \mathbb{R}^n$ and a real $n \times n$ orthogonal matrix Q so that $Q\mathbf{u} + \mathbf{p} = \alpha\mathbf{e}_n$, $Q\mathbf{v} + \mathbf{p} = \beta\mathbf{e}_n$, $Q\mathbf{w} + \mathbf{p} = \mathbf{0}$ where $\mathbf{e}_n$ is the nth coordinate vector in $\mathbb{R}^n$ and $\alpha < 0 < \beta$. Writing $Q\mathbf{h}(\mathbf{x}) + \mathbf{p}\,\|\mathbf{x}\|^2 = \widetilde{\mathbf{h}}(\mathbf{x})$, we see that for some $\mathbf{x},\mathbf{y} \in V_1$

$$\widetilde{h}_\ell(\mathbf{x}) = \widetilde{h}_\ell(\mathbf{y}) = 0, \quad \ell = 1,\dots,n-1$$

and

$$\widetilde{h}_n(\mathbf{x}) = \alpha, \quad \widetilde{h}_n(\mathbf{y}) = \beta.$$

It follows that the continuous form $\widetilde{h}_n$ takes both signs on Z_1, the joint zero set for $\widetilde{h}_1, \dots, \widetilde{h}_{n-1}$, and therefore $\widetilde{h}_n(\mathbf{z}) = 0$ for some $\mathbf{z} \in Z_1$. Thus $\widetilde{\mathbf{h}}(\mathbf{z}) = \mathbf{0} = Q\mathbf{w} + \mathbf{p}$, and hence $\mathbf{w} = \mathbf{h}(\mathbf{z})$, i.e. $\mathbf{w} \in W_1$. $\square$

We are now ready for convexity of joint ranges.

Theorem 4.2. *W_1 is convex if $n \leqslant d$, and if $n \leqslant d+1$ provided* $\dim V > 2$.

Proof. Suppose first that $\dim V < \infty$. Then our conclusions follow from Theorems 2.5, 3.3 and Lemmas 2.3(b) and 4.1. Now suppose that $\dim V = \infty$, and let $\mathbf{u}, \mathbf{v} \in W_1$.This means that $\mathbf{u} = \mathbf{h}(\mathbf{x})$, $\mathbf{v} = \mathbf{h}(\mathbf{y})$ for some $\mathbf{x}, \mathbf{y} \in V_1$. Choose a three dimensional subspace $L \subset V$ so that $\mathbf{x}, \mathbf{y} \in L$. By what has already been proved, the segment $t\mathbf{u} + (1-t)\mathbf{v}$, $0 < t < 1$, is contained in the joint range for $\mathbf{h}|_L$, and hence in the joint range W_1 for $\mathbf{h}$. $\square$

Corollary 4.3. *The results of Theorem* 4.2 *hold with W_1 replaced by W_+.*

Proof. Let $\mathbf{u} = \mathbf{h}(\mathbf{x})$, $\mathbf{v} = \mathbf{h}(\mathbf{y})$ with $\mathbf{x}, \mathbf{y} \in V_+$ and let $\mathbf{w} = t\mathbf{u} + (1-t)\mathbf{v}$, $0 < t < 1$. By (1.1), $\mathbf{u} = \|\mathbf{x}\|^2\mathbf{h}(\widehat{\mathbf{x}})$ and $\mathbf{v} = \|\mathbf{y}\|^2\mathbf{h}(\widehat{\mathbf{y}})$ where $\widehat{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$, $\widehat{\mathbf{y}} = \mathbf{y}/\|\mathbf{y}\| \in V_1$. Now write $\alpha = (t\|\mathbf{x}\|^2 + (1-t)\|\mathbf{y}\|^2)^{1/2}$ and $\beta = t\alpha^{-2}\|\mathbf{x}\|^2$. By Theorem 4.2, $\beta\mathbf{h}(\widehat{\mathbf{x}}) + (1-\beta)\mathbf{h}(\widehat{\mathbf{y}}) = \mathbf{h}(\widehat{\mathbf{z}})$ for some $\widehat{\mathbf{z}} \in V_1$. Thus $\mathbf{w} = \mathbf{h}(\alpha\widehat{\mathbf{z}}) \in W_+$. $\square$

5. Examples

We now give some examples illustrating the ideas above. The first illustrates the quaternionic numerical range for $V = \mathbb{H}$ with inner product $\langle p, q\rangle = p\bar{q}$.

Example 5.1. Let $A : V \to V : q \to q\mathrm{i}$, so its numerical range is

$$W(A) = \{q\mathrm{i}\bar{q} : q \in \mathbb{H}, |q| = 1\}.$$

If $|q| = 1$ then $|q\mathrm{i}\bar{q}| = 1$, so $W(A) \subset V_1$. Since $1\mathrm{i}\bar{1} = \mathrm{i}$, $\mathrm{j}\mathrm{i}\bar{\mathrm{j}} = -\mathrm{i}$ and $0 \notin V_1$, $W(A)$ is not convex.

Our second example illustrates the sharpness of Theorem 2.5.

Example 5.2. Let $V = \mathbb{H}^2$, $n = d = 4$; $\sigma_1(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_2 + x_2\bar{y}_1$, $\sigma_2(\mathbf{x}, \mathbf{y}) = x_2\bar{y}_1\mathrm{i} - \mathrm{i}x_1\bar{y}_2$, $\sigma_3(\mathbf{x}, \mathbf{y}) = x_2\bar{y}_1\mathrm{j} - \mathrm{j}x_1\bar{y}_2$, $\sigma_4(\mathbf{x}, \mathbf{y}) = x_2\bar{y}_1\mathrm{k} - \mathrm{k}x_1\bar{y}_2$, so $\mathbf{h}(\mathbf{x}) = 2(\mathrm{Re}(x_1\bar{x}_2), -\mathrm{Re}(\mathrm{i}x_1\bar{x}_2), -\mathrm{Re}(\mathrm{j}x_1\bar{x}_2), -\mathrm{Re}(\mathrm{k}x_1\bar{x}_2))$.

The corresponding quaternion $h(\mathbf{x}) = 2x_1\bar{x}_2$, so $Z_+ = \mathbf{h}^{-1}(\mathbf{0}) \cap V_+ = h^{-1}(0) \cap V_+ = \{\mathbf{x} \neq \mathbf{0} : x_1 = 0\} \cup \{\mathbf{x} \neq \mathbf{0} : x_2 = 0\}$ is disconnected.

We turn now to Theorem 4.2 for $\dim V = 2$.

Example 5.3. As for Example 5.2, but with $n = 5$ and $\sigma_5(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_1 - x_2\bar{y}_2$.

Thus $h_5(\mathbf{x}) = |x_1|^2 - |x_2|^2$ and $\|\mathbf{h}(\mathbf{x})\|^2 = |x_1|^2 + |x_2|^2 = \|\mathbf{x}\|^2$, so $\|\mathbf{h}(\mathbf{x})\| = 1$ for $\mathbf{x} \in V_1$, and hence $0 \notin W_1 = \mathbf{h}(V_1)$. Since $\mathbf{h}(1, 0) = \mathbf{e}_5$, and $\mathbf{h}(0, 1) = -\mathbf{e}_5$, W_1 is nonconvex.

To illustrate the sharpness of Theorem 3.3, we have

Example 5.4. Let $V = \mathbb{H}^3$, $n = 5$ and $\sigma_5(\mathbf{x}, \mathbf{y}) = x_3\bar{y}_3$ with σ_ℓ as in Example 5.2 for $\ell = 1, \ldots, 4$.

Then $Z_+ = \{\mathbf{x} \neq \mathbf{0} : x_1 = x_3 = 0\} \cup \{\mathbf{x} \neq \mathbf{0} : x_2 = x_3 = 0\}$ is disconnected.

Finally, we exhibit sharpness of Theorem 4.2 for $\dim V = 3$.

Example 5.5. Let $V = \mathbb{H}^3$, $n = 6$, σ_ℓ as in Example 5.3 for $\ell = 1, \ldots, 5$ and σ_6 as for σ_5 in Example 5.4.

Thus $\mathbf{h}(1, 0, 0) = \mathbf{e}_5, \mathbf{h}(0, 1, 0) = -\mathbf{e}_5$ and we claim that $\mathbf{0} \notin W_1$. Indeed if $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ then $h_6(\mathbf{x}) = |x_3|^2$ forces $x_3 = 0$, and $0 = \sum_{\ell=1}^{5} h_\ell(\mathbf{x})^2 = |x_1|^2 + |x_2|^2$ shows that $\mathbf{x} = \mathbf{0}$. Thus W_1 is not convex.

Remark 5.6

(a) By Lemma 4.1, optimality of Theorems 2.5 and 3.3 follows from Examples 5.3 and 5.5, but we felt it worthwhile to give the explicit Examples 5.2 and 5.4.

(b) It can be shown that in Example 5.3, $W_1 = S^4$ (the unit sphere in $\mathbb{R}^5$), and in Example 5.5, $W_1 = S^5$. In particular, if we replace h_5 in Example 5.3 by $h_5 + 1$ then W_1 becomes a (nonconvex) sphere of radius one centred at $\mathbf{e}_5$, but W_+ becomes a (convex) half space $\{\mathbf{x} : x_5 > 0\} \cup \{\mathbf{0}\}$.

(c) Examples similar to 5.4 and 5.5 may be easily constructed for $\mathbb{H}^m$, $3 < m \leqslant \infty$, where we define $\mathbb{H}^\infty = \ell^2_{\mathbb{H}} = \left\{\mathbf{x} = (x_j)_1^\infty : x_j \in \mathbb{H}, \sum_1^\infty |x_j|^2 < \infty\right\}$.

(d) Examples similar to 5.2–5.5 may be easily constructed for $\mathbb{F} = \mathbb{C}$, and they are well-known.

Acknowledgements

P. Binding's research was supported by NSERC of Canada. A. Markus thanks the University of Calgary for support and hospitality.

References

[1] F.V. Atkinson, Multiparameter Eigenvalue Problems, vol. 1, Academic Press, New York, 1971.

[2] Y.-H. Au-Yeung, A simple proof of the convexity of the field of values defined by two Hermitian forms, Aequations Math. 12 (1975) 82–83.
[3] Y.-H. Au-Yeung, On the convexity of numerical range in quaternionic Hilbert space, Linear Multilinear Algebra 16 (1984) 93–100.
[4] Y.-H. Au-Yeung, Y.-T. Poon, A remark on the convexity and positive definiteness concerning Hermitian matrices, Southeast Asian Bull. Math. 3 (1979) 85–92.
[5] Y.-H. Au-Yeung, N.-K. Tsing, An extension of the Hausdorff–Toeplitz theorem on the numerical range, Proc. Amer. Math. Soc. 89 (1983) 215–218.
[6] P. Binding, Hermitian forms and fibration of spheres, Proc. Amer. Math. Soc. 94 (1985) 581–584.
[7] P. Binding, P.J. Browne, A variational approach to multiparameter eigenvalue problems for matrices, SIAM J. Math. Anal. 8 (1977) 763–777.
[8] L. Brickman, On the field of values of a matrix, Proc. Amer. Math. Soc. 12 (1961) 61–66.
[9] M. Fan, A. Tits, m-form numerical range and the computation of the structured singular value, IEEE Trans. Autom. Control 33 (1988) 284–289.
[10] A. Feintuch, A. Markus, The Toeplitz–Hausdorff theorem and robust stability theory, Math. Intell. 21 (3) (1999) 33–36.
[11] A. Feintuch, A. Markus, The structured norm of a Hilbert space operator with respect to a given algebra of operators, Operat. Theor.: Adv. Appl. 115 (2000) 163–183.
[12] D.W. Fox, The convexity of the range of three Hermitian forms and of the numerical range of sesquilinear forms, Linear Algebra Appl. 22 (1978) 191–194.
[13] F. Hausdorff, Der Wertvorrat einer Bilinearform, Math. Z. 3 (1919) 314–316.
[14] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
[15] V.I. Istratescu, Inner Product Structures, Reidel, Dordrecht, 1987.
[16] Yu. Lyubich, Separation of roots of matrix and operator polynomials, Integral Equations Operator Theory 29 (1997) 52–62.
[17] Yu. Lyubich, A. Markus, Connectivity of level sets of quadratic forms and Hausdorff–Toeplitz type theorems, Positivity 1 (1997) 239–254.
[18] A. Markus, L. Rodman, Some results on numerical ranges and factorisations of matrix polynomials, Linear Multilinear Algebra 42 (1997) 169–185.
[19] W. So, An explicit criterion for the convexity of quaternionic numerical range, Canad. Math. Bull. 41 (1998) 105–108.
[20] O. Toeplitz, Das algebraische Analogen zu einem Satze von Fejér, Math. Z. 2 (1918) 187–197.